مجلة العلوم البحثة والتطبيقية
Journal of Pure & Applied Sciences
www.Suj.sebhau.edu.ly ISSN 2521-9200
Received 05/01/2018 Revised 02/02/2018 Published online 28/03/2018

A study on max domain of attraction condition of Fréchet distribution with case study

H. A. Alaswed
Statistics Department, Faculty of Science, Sebha University, Libya
Corresponding Author: haf.alaswed@sebhau.edu.ly

Abstract Extreme value theory (EVT) refers to branch of statistics that deals with extreme events (minimum or maximum). In EVT, Fréchet distribution is one of the most important to model extreme events. In this paper, the first objective is to discuss the problem which is called the domain of attraction condition of maximum belonging to the domain of Fréchet distribution by using Gnedenko's, von Mises, Castillo and de Hann necessary and sufficient condition. The second objective, graphical methods are used to detect an appropriate distribution. The third objective of the present study is to model the behaviour of annual maximum wind speed data at Albany, and Hartford by using Fréchet distribution. The distribution parameters will be estimated by Maximum Likelihood Estimation (MLE). The results show that the two applications are significant to be fitted by Fréchet model. All computations in this work are performed by R programming with packages of fExtreme and ismev for parameter estimation and diagnostic plots.

Keywords: Domain of Attraction Condition, Fréchet Distribution, Norming Constant, P-P Plot, Q-Q Plot, Return Level Plot.

دراسة شروط مجال الجاذبية للقيم العظمي لتوزيع فريشت مع دراسة حالة

حافظ ابوبكر امحمد
قسم الإحصاء– كلية العلوم–جامعة سبها، ليبيا
للمراسلة:haf.alaswed@sebhau.edu.ly

الملخص تستند نظرية القيمة المتطرفة على تحليل القيم الصغرى والعظمي لدراسة بعض الظواهر. في نظرية القيم المتطرفة يعتبر توزيع فريشيت أحد التوزيعات الهامة حيث يستخدم لنمذجة الأحداث المتطرفة. الهدف الأول في هذه الورقة، هو مناقشة مشكلة شروط مجال جذب توزيعات القيم العظمي لتنتمي إلى مجال توزيع فريشيت باستخدام شروط كلا من جينيدينكو، فون ميسس، كاستيلو ودي هان. الهدف الثاني هو استخدام رسومات بيانية للكشف عن التوزيع المناسب. والهدف الثالث في هذه الدراسة هو نمذجة سلوك سرعة الرياح باستخدام توزيع فريشيت. تم استخدام طريقة الإمكان الأعظم لتقدير معلمات التوزيع وأظهرت النتائج أنه في جميع الحالات من التطبيقين أن نموذج فريشيت مناسبا. حيث تم استخدام برنامج R في العمليات الحسابية لتقدير المعالم وإجراء الرسومات البيانية التشخيصية.

الكلمات المفتاحية: توزيع فريشت، ثوابت الاتزان، رسمة الاحتمال، رسمة التجزئة، رسمة مستوي العائد، شروط مجال الجاذبية.

1 Introduction

Extreme value theory (EVT) has emerged as one of the most important statistical disciplines for the applied sciences and widely used in many other disciplines over the last 50 years. Extreme events are identified with the probability of the events as more extreme than any that already been observed. The purpose of extreme value analysis (EVA) is to model the rare events by finding reliable estimates of the frequency of these events, see,[1]. A theoretical definition and more detailed explanation can be found in, [2]. The asymptotic theory of extreme value distributions were introduced in, [3] . Some useful sufficient conditions have been presented by Mises, [4], for the weak convergence of the three types of limit distributions. We shall discuss von Mises conditions in a subsequent section. Gnedenko presents a rigorous foundation for the EVT and provided necessary and sufficient conditions for the weak convergence of the extreme order statistics in, [5]. His work was refined by, [6], who gives the most complete and rigorous discussion of the necessary and sufficient conditions problem.

This paper is organized as follows: Section 2, describes a brief review of some theoretical results concerning on maximum Fréchet distribution. In section 3, some graphical method diagnostic of P-P plot, Q-Q plot and R-L plot are presented. Maximum domain of attraction condition and norming constants of Fréchet distribution are given in section 4. In this section, the condition of Gnedenko's, von Mises, Castillo and de Hann are presented with some examples. In section 5, the case study of annual maximum of wind speed to model the behaviour of extreme by using Fréchet distribution is described. Finally, concluding remarks and future work is given in section 6.

2 Fréchet distribution

In EVT, Fréchet distribution is one of the probability distributions of extreme value and widely used in several areas for modelling of extreme events such as: flood, rainfall, temperature and wind speeds etc.[7]. Fréchet distribution is a special case of the generalized extreme value distribution (GEVD) when shape parameter $\gamma > 0$ and also known as inverse Weibull distribution. The probability density function (pdf) of maximum Fréchet ($F_{\max}$) is given by:

$$f(x) = \frac{\gamma\sigma}{(x-\mu)^2}\exp[-(\frac{\sigma}{x-\mu})^{\gamma}](\frac{\sigma}{(x-\mu)})^{\gamma-1}, x > \mu \quad \ldots 1$$

And has the cumulative distribution function (cdf):

$$F_M(x;\gamma,\mu,\sigma) = \begin{cases} 0 & , x < \mu, \\ \exp\left\{-\left(\frac{\sigma}{x-\mu}\right)^{\gamma}\right\} & , x \ge \mu, \gamma > 0 \end{cases} \quad ..2$$

Where γ,μ,σ are shape, location and scale parameter respectively. The shape parameter $\gamma = 1/\alpha$ is called the extreme value index (EVI). Four important functions associated with random variables of Fréchet distribution in minimum and maximum cases, The survival ($\bar{F}(x)$), hazard ($H(x)$) characteristic and quantile functions are presented in[8] and [10].

3 Graphical Method Diagnostic

In this section, we present some graphical plots to select a suitable to make conclusions about some aspect of the population from which the data were drawn. Such conclusions can be sensitive to the accuracy of the fitted model, so it is necessary to check that the model fits well. One of the most commonly used graphical methods in,[9], statistics to build several types of graphical and how they can be used to select a parent distribution for the given sample. Now, we give some graphical methods for checking whether a fitted model is in agreement with the data such that, probability -probability (P-P) plot, Quantile Quantile (Q-Q) plot and Return level (R-L) plot by [9], that deals with the problem of selection models by diagnostic plots. Firstly, we describe the P-P Plot.

3.1 The P-P Plot

The P-P Plot can be used to distinguish visually between different distribution. Let $x_1, x_2, \ldots, x_n$ be a sample from a given population with estimated cdf $\hat{F}(x)$. The scatter plot points of P-P plot are given by:

$$\hat{F}(x_{i:n}) \text{ versus } P_{i:n}, i = 1, 2, \ldots, n \quad \ldots 3$$

Where $P_{i:n}$ is called plotting positions (PP) is defined by :

$$P_{i:n} = \frac{i-\alpha}{n+\beta}, \quad i = 1, 2, \ldots, n \quad \ldots 4$$

For appropriate choices of $\alpha, \beta \ge 0$. In PP α and β values can be chosen empirically (depending on the data, the type of distribution, the estimation method to be used, etc.). Here we use $\alpha = 0$ and $\beta = 1$ that is,

$$P_{i:n} = \frac{i}{n+1}, \quad i = 1, 2, \ldots, n \quad \ldots 5$$

Other alternative can be found in [10]. If the model fits the data well, the points on the graph would lie along a positively straight line. Substantial departures from linearity provide evidence of a failure in $\hat{F}$ as a model for the data. For more details on diagnostic plots, see,[10].

3.2 The Q-Q Plots

The quantile-quantile plot (Q-Q plot) is a convenient visual tool to examine whether a sample comes from a specific distribution. Let $\hat{F}(x)$ be an estimate of $F(x)$ based on $x_1, x_2, \ldots, x_n$. The scatter plot of the points:

$$\hat{F}^{-1}(P_{i:n}) \text{ versus } x_{i:n}, i = 1, 2, \ldots, n \quad \ldots 7$$

is called a Q-Q plot, thus, the Q-Q plot show the estimated versus the observed quantiles. If the model fits the data well, the pattern of points on the Q-Q plot will exhibit a straight line and the picture strongly suggests that the data follow a distribution is acceptable.

3.3 Return Level Plot (RLP)

From the fitted Fréchet distribution, we can estimate how often the extreme quantiles occur with a certain return level. The return value is defined as a value that is expected to be equalled or exceeded on average once every interval of time (T) (with a probability of 1/T). In [2], the return value can be calculated by solving this equation (i.e., by inverting Eq-2 in Max-Fréchet). Quantile Max-Fréchet is:

$$X_p = \mu + \sigma(-\log p)^{-1/\gamma} \quad \ldots 8$$

We estimate above parameters by using MLE and substituting the parameters γ,μ,σ by their estimates $\hat{\gamma},\hat{\mu},\hat{\sigma}$ to get $\hat{x}_p$. The RLP represents the points $(\log y_p, x_p)$ where $y_p = -\log p$ and confidence intervals are usually added to this plot to increase its information. Consequently, on max the domain attraction condition of Fréchet distributions appear as concave.

4 Domain Attraction of Fréchet Distribution (DAFD)

In this section, we shall describe the necessary and sufficient conditions for a distribution function F to belong to the Max-domain attraction of $F_{\max}$. Before that, we will introduce some basic concept that provide most of the background necessary to fully understand the mathematical derivations for the distribution to belong to Max-domain attraction of Fréchet distribution.

4.1 Some basic concept

Let F be a distribution function and $F^{-1}(y) = \inf\{x : F(x) \succ y\}$ is an inverse function of F. Let $x^r = \sup\{x : F(x) \prec 1\}, x^L = \inf\{x : F(x) > 0\}$ the right and left end point respectively. We define the tail quantile function U by $U(t) = F^{-1}(1-\frac{1}{t})$ or equivalently $U(t) = (1/1-F)^{-1}(t)$. The relationship between the tail quantile function and the left and right end point written as follows:

$$x^L = inf\{x : F(x) > 0\} = U(1) = F^{-1}(0)$$

$$x^r = \sup\{x : F(x) \prec 1)\} = U(t=\infty) = F^{-1}(1-\frac{1}{t}) = F^{-1}(1)$$

4.2 Maximum Domain of Attraction (MDA, $F \in D_{\max}(F_M)$)

We now present a detailed discussion for cd F to be in the domain attraction of Max-Fréchet distribution by introducing in the following four definitions :

Definition-1: **Gnedenko's necessary and sufficient condition for MDA**

A distribution function F is said to be in the Max-domain of Fréchet distribution and we write $F \in D_{\max}(F_M)$ if and only if

$x^r = F^{-1}(1) = \infty$ and

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-\alpha}, x > 0, \alpha = \frac{1}{\gamma} \qquad \ldots 9$$

Where x^r is the right end point and F^{-1} is generalized inverse function F .

In [11], Gnedenko's necessary and sufficient conditions for $F \in D_{\max}(F_M)$ as given in Eq-9 are quite often difficult to verify. For this reason, von Mises established useful sufficient conditions in [4], and very easy to check, requiring only the existence of the first or second derivative of F , but are only applicable to absolutely continuous d.f.'s F . Besides that, they are only sufficient conditions and not necessary. The sufficient conditions of von Mises are widely studied in this day and easily applicable can be definition as follows:

Definition-2: Von Mises sufficient condition for MDA:

Let F be a distribution function and x^r its right endpoint. Suppose F' and F'' are exists and let $H(x)$ hazard function which we have already seen in [10]. The distribution function F is in Max-domain attraction of Fréchet and we write $F \in D_{\max}(F_M)$ if and only if

$$\lim_{x\to\infty} xH(x) = \lim_{x\to\infty} \frac{xf(x)}{1-F(x)} = \alpha \qquad \ldots 10$$

An alternative formulation of sufficient condition is given by, [10], in the following definition.

Definition-3: Castillo and Hadi of sufficient condition for MDA

The distribution function F is in the Max-domain attraction of Fréchet distribution and we write $F \in D_{\max}(F_M)$ if and only if

$$\lim_{e\to\infty} \frac{F^{-1}(1-e) - F^{-1}(1-2e)}{F^{-1}(1-2e) - F^{-1}(1-4e)} = 2^{-\gamma}, \gamma < 0 \qquad \ldots 11$$

Where e is the base of the natural logarithm and γ is the shape parameter. In [6], it shows that the last one of necessary and sufficient conditions in term of U have been given in the tail quantile function of a cdf of F .

Definition-4: de Haan and Ferreira of sufficient conditions for MDA

A cd F is in the Max-domain attraction of Fréchet distribution with $\gamma > 0$ and we write $F \in D_{\max}(F_M)$ if and only if

$$\lim_{t\to\infty} \frac{tU'(t)}{U(t)} = \gamma \qquad \ldots 12$$

or equivalent to

$$\lim_{t\to\infty} \frac{tU(tx)}{U(t)} = x^{\gamma} \qquad \ldots 13$$

Finally, the possible choices for norming constants, scale a_n and location b_n , these constants are not unique, depend on the type of domain attraction. We can choose the location and scale in Max-domain of Fréchet as follows:

4.3 Norming constant of Max-domain of Fréchet

$$scale = a_n = \inf\{x : 1 - F(x) \prec F^{-1}(1-\frac{1}{n})$$

and $Location = b_n = 0$ ... 14

For more details about the norming constants scale and location, see,[11].

4.4 Some Examples: In this section, we will give a number of examples to understand the Max-domain of Fréchet distribution, that illustrate in [8], how to find the domain of attraction. We want to introduce four examples, the first two examples do not belong and the other two examples belong to Fréchet distribution. We begin with a simple example that can easily be generalized.

Example-1 : Let X be a random variable with uniform distribution $U(0,1)$ has cdf: $F(X) = x \;\; ; 0 \le x \le 1$. Since $x^r = F^{-1}(1) = 1$ and also $x^L = F^{-1}(0) = 0$ fails to hold. Thus, the uniform distribution does not belong to domain of Fréchet distribution and we write $F \notin D_{\max}(F_M)$.

Example-2: Let $F(x) = 1 - (\log x)^{-1}, x \ge e$. Since $x^r = +\infty$ satisfies and next condition we are going to determine the sufficient condition of Genedenko

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(x)} = \lim_{t\to\infty} \frac{1-[1-(\log tx)^{-1}]}{1-[1-(\log t)^{-1}]} =$$

$$\lim_{t\to\infty} \frac{(\log tx)^{-1}}{(\log t)^{-1}} = \lim_{t\to\infty} \frac{\log t}{\log tx} = \lim_{t\to\infty} \frac{1}{t} \frac{1}{x\frac{1}{tx}} = 1 \ne x^{-\infty}$$

Gnedenko's sufficient conditions it fails to hold. Thus, $F \notin D_{\max}(F_M)$.

Example-3: Suppose F has a Pareto distribution with cdf $F(x) = 1 - x^{-\alpha}, x \ge 1, \alpha > 0$. For determining the Max-domain attraction of Fréchet. Since $x^r = +\infty$, then we need to determine the sufficient condition of Genedenks:

$$\lim_{t\to\infty}\frac{1-F(tx)}{1-F(t)}=\lim_{t\to\infty}\frac{1-[1-(tx)^{-\theta}]}{1-[1-t^{-\theta}]}=$$

$$\frac{t^{-\theta}x^{-\theta}}{t^{-\theta}}=\lim_{t\to\infty}x^{-\theta}=x^{-\theta}$$

In this case we say that F is in the Max-domain attraction of Fréchet distribution and the notation $F\in D_{\max}(F_M)$. When the domain of attraction condition is satisfied; we can choose the normalizing constants of Max-domain of Fréchet distribution by the expressions:

$$b_n=0,\qquad a_n=F^{-1}\left(1-\frac{1}{n}\right)=\left(1-1+\frac{1}{n}\right)^{-\frac{1}{\theta}}=\left(\frac{1}{n}\right)^{-\frac{1}{\theta}}$$

Also we will try to check whether the von Mises sufficient condition (Eq-10) holds. For that purpose, let us note that:

$$\lim_{t\to\infty}\frac{xf\left(x\right)}{1-F\left(x\right)}=\lim_{t\to\infty}\frac{\theta x^{-\theta}}{x^{-\theta}}=\theta$$

This means that the condition of von Mises hold and we conclude that $F\in D_{\max}(F_M)$

Example-4: Let $F(x)=\frac{1}{2}+\frac{1}{\pi}\tan^{-1}x,-\infty<x<\infty$

Since $x^{r}=+\infty$ hold, then we will try to check the sufficient condition (Eq-10) holds. For that purpose, we find:

$$\lim_{t\to\infty}\frac{1-F(tx)}{1-F(t)}=\lim_{t\to\infty}\frac{xf(tx)}{f(t)}=$$

$$\lim_{t\to\infty}\frac{x(2t)}{(2tx^{2})}=\lim_{t\to\infty}\frac{x}{x^{2}}=\frac{1}{x}=x^{-1}=x^{-\alpha}$$

This means that the condition of Max-domain of Fréchet hold, which shows that the Cauchy distribution belongs to the Max-domain attraction of Fréchet and we write $F\in D_{\max}(F_M)$. Finally, the possible choices for the norming constants of this distribution is:

$$a_n=F^{-1}\left(1-\frac{1}{n}\right)=\tan\left(\pi\left(1-\frac{1}{n}\right)-\frac{\pi}{2}\right)=\qquad b_n=0$$

$$\tan\left(\pi-\frac{\pi}{n}-\frac{\pi}{2}\right)=\tan\left(\frac{\pi}{n}-\frac{\pi}{2}\right)$$

Some examples of distributions which belong to the domain attraction of Fréchet distributions are given in [8]. We note that the domain attraction condition of Gnedenko, von Mises, Castillo and de Haan can easily be used to derive explicit examples in the Max-domain attraction condition of Fréchet distributions.

5 Case study

The data sets used in this study consist of annual maximum (AM) to climate data of wind speed at Albany, and Hartford over the period 1944 to 1983. The source data is taken from ismev package. **Fig. 1** displays the AM data of wind speed at Albany, and Hartford. On the left panel of **Fig. 1** plot the AM wind speed data of Albany and the right panel of Hartford.

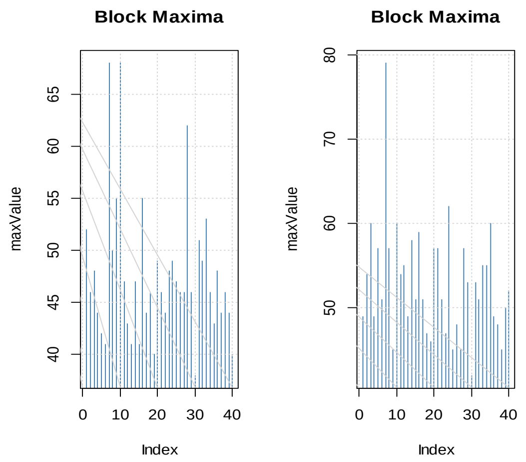

Fig. 1: Annual maximum data of wind speed at Albany (left), and Hartford (right).

5.1 Summary Statistics

Firstly, we present some results based on descriptive statistics of wind speed application data of two countries and the results are reported in Table-1.

Table 1: Descriptive statistics of wind data

	Country	
Statistics	Albany	Hartford
Min	38	42
Max	68	79
Median	46	51.5
Mean	47.58	52.83
Skewness	1.53	1.40

In Table 1, it can be seen that the results of all descriptive statistics at Albany are smaller than Hartford except skewness. The skewness are positive for two country, that mean the distribution has right tail, so there is a good reasons to think that the distribution of these data belong to the Max-domain of Fréchet distribution ($\gamma>0$). We apply MLE to estimate the three parameters of distribution for the application data. The AM data is fitted by MLE to get the point estimate and summarizes the results in Table-2.

Table 2: Parameters estimation of Fréchet distribution by MLE.

Parameters	Methods	Type of data	
	S.D	Albany	Hartford
Shape	MLE	0.09	0.03
Location	MLE	44.58	49.93
Scale	MLE	4.36	5.01
Tail behaviour		$\gamma>0$	$\gamma>0$

The results listed in Table 2, show the estimates of three parameters. The point estimates of shape for two types of data are positive and indicate that the distribution of data has right tail and Fréchet distribution is an appropriate for these data of wind speed.

5.2 Diagnostic Plots

In order to get an idea about the tail behaviour of the distribution. We present in **Fig 2-3**, the various diagnostic plots for assessing the accuracy of the model fitted for an application data of AM observations of wind speed at Albany, and Hartford are shown in four different plots; P-P-plot, Q-Q-plot, Return level plot and density plot.

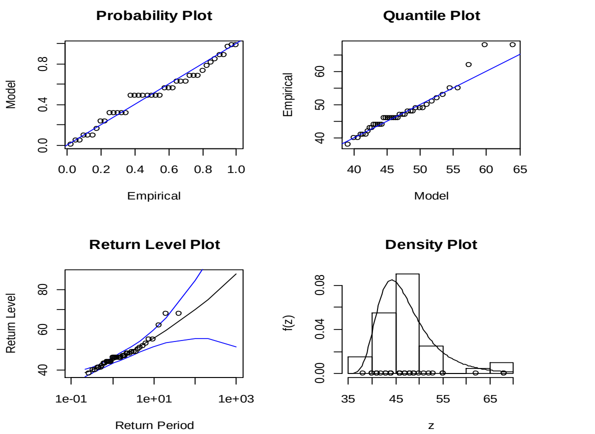

Fig. 2: Four different plots; P-P, Q-Q, R-L and density plot of Albany.

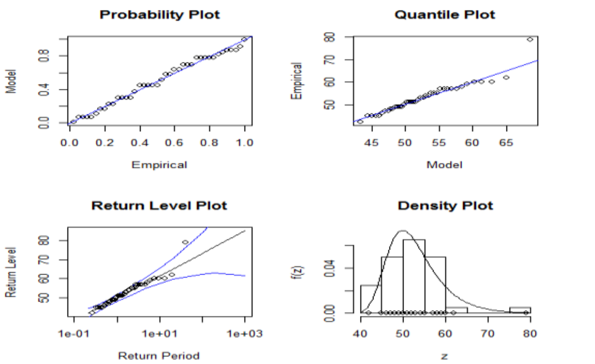

Fig. 3: Four different plots; P-P, Q-Q, R-L and density plot data of Hartford.

The graph of the P-P and the Q-Q plot are shown the data came from a Fréchet distribution. the R-L plot suggest that the model departures of the positive shape estimate. Finally, the corresponding density estimate seems consistent with the histogram of the data. Consequently, all four diagnostic plots lend support to Fréchet model are the best fit.

6 Conclusion

Extreme Value Theory (EVT) is a progressive branch of statistics dealing with extreme events. In EVT, Fréchet distribution has wide application in modelling extreme events. In this paper, we discuss the problem which is called the Max-domain of attraction condition for belonging to the domain of Max-Fréchet distribution by using Gnedenko's, von Mises, Castillo and de Hann necessary and sufficient condition and practical aspects of the use Fréchet distribution. We also studied some graphical methods such as: Q-Q plots, P-P plot and R-L plots. To illustrate how extreme value theory can be used to model extreme events, when we used the sample data of wind speed. For each country, the Q-Q plot, P-P plot and R-L Plot shows that the scatter of points are supported the assumption of Fréchet distribution model is useful to describe the behavior of this kind of data, also the corresponding density estimate seems consistent with the histogram of the data. Some issues that would be considered for future works in this study are the problem of Min-domain of attraction condition of Fréchet distribution and other some diagnostics tests in order to select the best fit model. Finally, all results of problem in the domain of attraction condition of Fréchet distribution (under linear) will try reformulated under power and we look forward to see other application to recognize.

References

[1]- X. Liu, “Data analysis in extreme value theory,” *Dep. Stat. Oper. Res. Univ. North Carolina Chapel Hill*, 2009.

[2]- S. Coles, J. Bawa, L. Trenner, and P. Dorazio, *An introduction to statistical modeling of extreme values*, vol. 208. Springer, 2001:P37,P49.

[3]- R. A. Fisher and L. H. C. Tippett, “Limiting forms of the frequency distribution of the largest or smallest member of a sample,” in *Mathematical Proceedings of the Cambridge Philosophical Society*, 1928, vol. 24, no. 2, pp. 180–190.

[4]- R. Von Mises, “Über allgemeine Quadraturformeln.,” *J. für die reine und Angew. Math.*, vol. 174, pp. 56–67, 1936.

[5]- B. Gnedenko, “Sur la distribution limite du terme maximum d’une serie aleatoire,” *Ann. Math.*, pp. 423–453, 1943.

[6]- L. De Haan and A. Ferreira, *Extreme value theory: an introduction*. Springer Science & Business Media, 2007:PP17-21.

[7]- K. Abbas and T. Yincai, “Comparison of Estimation Methods for Frechet Distribution with Known Shape.,” *Casp. J. Appl. Sci. Res.*, vol. 1, no. 10, 2012.

[8]- J. Beirlant, Y. Goegebeur, J. Segers, and J. L. Teugels, *Statistics of extremes: theory and applications*. John Wiley & Sons, 2006:PP56-59.

[9]- E. Omey, F. Mallor, and E. Nualart, “An introduction to statistical modelling of extreme values. Application to calculate extreme wind speeds,” 2009.

[10]- E. Castillo, A. S. Hadi, N. Balakrishnan, and J.-M. Sarabia, *Extreme value and related models with applications in engineering and science*. Wiley Hoboken, NJ, 2005:P108,P148

[11]- S. Kotz and S. Nadarajah, *Extreme value distributions: theory and applications*. World Scientific, 2000:PP210-213